Available online at www.sciencedirect.com
ScienceDirect
Journal of Pharmaceutical and Biomedical Analysis 44 (2007) 683–700
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Review

A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies

Yves Roggo *, Pascal Chalus, Lene Maurer, Carmen Lema-Martinez, Aurélie Edmond, Nadine Jent

F. Hoffmann-La Roche Ltd., Basel, Switzerland

Received 24 January 2007; received in revised form 15 March 2007; accepted 21 March 2007
Available online 30 March 2007

Abstract

Near-infrared spectroscopy (NIRS) is a fast and non-destructive analytical method. Associated with chemometrics, it becomes a powerful tool for the pharmaceutical industry. Indeed, NIRS is suitable for analysis of solid, liquid and biotechnological pharmaceutical forms. Moreover, NIRS can be implemented during pharmaceutical development, in production for process monitoring or in quality control laboratories.This review focuses on chemometric techniques and pharmaceutical NIRS applications. The following topics are covered: qualitative analyses, quantitative methods and on-line applications. Theoretical and practical aspects are described with pharmaceutical examples of NIRS applications.

Keywords: Near infrared spectroscopy; Chemometrics; Pharmaceuticals; On-line; Quality control; Identification; Qualification; Quantification

Contents

Abbreviations: ANN, artificial neural networks; API, active pharmaceutical ingredient; ATR, attenuated total reflectance; CC, correlation coefficient; CFR, code of federal regulations; DBSCAN, density based spatial clustering of applications with noise; DS, direct standardization; DSC, differential scanning calorimetry; EMEA, European agency for the evaluation of medicinal products; FDA, food and drug administration; FT-IR, Fourier transform infrared spectroscopy; FT-NIR, Fourier transform near infrared spectroscopy; GC, gas chromatography; GMP, good manufacturing practice; HPLC, high performance liquid chromatography; ICH, international conference on harmonization; KF, Karl Fischer; KNN, K nearest neighbours; LDA, linear discriminant analysis; LOD, loss on drying; LVQ, learning vector quantization; MLR, multi-linear regression; MPE, mean percent error; MSC, multiplicative scatter correction; NIR, near infrared; NIRS, near infrared spectroscopy; NMR, nuclear magnetic resonance; OSC, orthogonal signal correction; PASG, pharmaceutical analytical sciences group; PC, principal component; PCA, principal component analysis; PCR, principal component regression; PDS, piecewise direct standardization; PLS, partial least squares regression; PLS-DA, partial least squares discriminant analysis; PNN, probabilistic neural network; QDA, quadratic discriminant analysis; RMSEC, root mean square error of calibration; RMSECV, root mean square error of cross-validation; RMSEP, root mean square error of prediction; RNA, ribonucleic acid; SEC, standard error of calibration; SECV, standard error of cross-validation; SEP, standard error of prediction; SIMCA, soft independent modelling of class analogy; SNV, standard normal variate; SVM, support vector machines; SVR, support vector regression; SW, Shenk and Westerhaus method; TG, thermogravimetry; XRD, X-ray powder diffractometry

* Corresponding author at: Process Robustness Support, Building 65, Room 516, F. Hoffmann-La Roche Ltd., Grenzacherstrasse, CH 4070 Basel, Switzerland. Tel.: +41 61 68 81 336.
E-mail address: yves.roggo@roche.com (Y. Roggo).

0731-7085/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2007.03.023

1. Introduction

In the 1800s William Herschel had discovered radiation beyond the visible red light. However, prior to Second World War the near infrared (NIR) region was not considered useful for spectroscopy [1]. It was observed that near infrared bands are severely overlapping and difficult to interpret.

Near Infrared Spectroscopy (NIRS) covers the transition from the visible spectral range to the mid-infrared region. In the area of NIR (800–2500 nm, respectively 12821–4000 cm^{-1}) [2] mainly vibrations of –CH, –OH, –SH and –NH bonds are observed. All the absorption bands are the results of overtones or combinations of the fundamental mid-infrared bands [3]. Many handbooks and papers describe the theory of near infrared spectroscopy [2,4].

Nowadays near infrared spectroscopy and chemometrics have proven their effectiveness for both qualitative and quantitative analyses in as different fields as agriculture [1,5], food [6], chemical [7] and oil industry [8].

NIRS is generally chosen for its speed, its low cost and its non-destructive characteristic towards the analyzed sample. On one hand, the interest in NIR has increased thanks to the instrument improvements and the development of fibre optics that allow the delocalization of the measurements. On the other hand it has increased because of the computer progresses and the development of new mathematical methods allowing data treatment. While mid-IR spectra and especially the absorbance bands are directly interpretable due to chemical peak specificity, NIR spectra are difficult to interpret.

Therefore, the use of chemometrics is required. Chemometrics [9–11] is a discipline using mathematical and statistical methods for the selection of the optimal experimental procedure and data treatment of chemical analyses. Chemometrics regroups several topics such as design of experiments, information extraction methods (modelling, classification and test of assumptions) and techniques allowing understanding the chemical mechanisms. A review concerning chemometrics has been written by Lavine [12] and many textbooks are available [13–15].

In this paper, the commonly used chemometric methods for the analysis of NIR spectra will be described. The three main techniques groups are the following:

- **Mathematical pretreatments** to enhance the information that is searched for the study, and decrease the influence of the side information contained in the spectra. The spectral pre-processing is considered as well known and not described in this text. The classical pretreatments are normalizations, derivatives and smoothing. For more details, readers are referred to textbooks [15,16].
- **Classification methods** to group samples together according to their spectra (description in the second part of the paper).
- **Regression methods** to link the spectrum to quantifiable properties of the samples (description in third part of the paper).

Near infrared spectroscopy is described in the European Pharmacopoeia (chapter 2.2.40) and in the US Pharmacopoeia general

chapter ⟨1 1 1 9⟩. Moreover guidelines are published by the European Agency for the Evaluation of Medicinal Products (EMEA/CVMP/961/01) and by the Pharmaceutical Analytical Sciences Group (PASG). We also want to mention previous reviews concerning the near infrared spectroscopy [17–20].

The aim of this article is to carry out a review concerning the pharmaceutical applications of near infrared spectroscopy and chemometrics. The first part focuses on the methods for qualitative analyses and classifications. In the second part regression methods and quantitative applications of NIRS will be described. In the last part examples of on-line applications will be studied.

2. Qualitative analyses by near infrared spectroscopy

Qualitative analysis means classification of samples according to their NIR spectra. NIR identifications are based on pattern recognition methods. Application of pattern recognition methodologies within chemistry [21], biology [22], and food sciences [23] are important.

The classification techniques can be divided into two categories: the unsupervised and the supervised ones. Concerning the unsupervised classification, samples are classified without a prior knowledge, except the spectra. Then the spectroscopist needs to explain these clusters. Supervised pattern recognition are techniques in which a prior knowledge, i.e., the category membership of samples, is required. Thus, the classification model is developed on a training set of samples with known categories [15]. Then the model performance is evaluated by comparing the classification predictions to the true categories of the validation samples.

The aim of this part is to describe briefly the chemometric methods for classification and to present an overview of pharmaceutical applications in the field of qualitative analyses, especially identification and qualification of raw and final materials.

2.1. Chemometric methods

2.1.1. Unsupervised classification methods

Principal component analysis (PCA) is a feature reduction method which forms the basis for multivariate data treatment. PCA is used to visualize the data. The most important PCA application is the reduction of the number of variables (scores) and the representation of a multivariate data table in a low dimensional space [16,24]. Thus, the new variables (loadings) are linear combinations of the original ones and can be interpreted like spectra.

Samples having different origins were analyzed in an article from Roggo et al. [25]. The aim of this study was to propose solutions to understand the differences between production sites. PCA was computed, the score plot confirmed statistical differences between the production sites, and the loadings identified the key-wavelengths and showed that the excipients were responsible of the differences.

Concerning the clustering methods, on one hand, hierarchical methods proceed by a successive divisions of the data set and result in a cluster sequence which can be represented with a tree, i.e., a dendrogramme [26]. On the other hand, there are non-hierarchical methods like Gaussian mixture models [26], K-means [10,27], density based spatial clustering of applications with noise (DBSCAN) [28,29] or Kohonen neural network [30]. Applications of unsupervised methods will be described in Section 2.2. The unsupervised methods, even if used in pharmaceutical environment, are not the most common ones. It is more likely to find supervised methods, which are going to be detailed in the very next section.

2.1.2. Supervised classification methods

There are three major differences between supervised pattern recognition algorithms [31]. There is a first distinction between methods focusing on discrimination, such as linear discriminant analysis (LDA), and those which put the emphasis more on similarity within a class, for example soft independent modelling of class analogy (SIMCA). The second difference concerns linear and non-linear methods like neural methods. The third distinction divides the parametric and non-parametric computations. In the parametric techniques such as LDA, statistical parameters of the normal distribution of samples are used in the decision rules.

The classical methods for the supervised classification are correlation based methods, distance based methods, LDA, SIMCA, and partial least squares discriminant analysis (PLS-DA).

2.1.2.1. Correlation and distances based methods. To be able to cluster objects, their similarity or dissimilarity is measured. Common mathematical methods in NIRS are correlation coefficient (CC) and distances in order to express similarity [32]. CC is defined as the cosine of the angle between the vector for the sample spectrum and the one for the average spectrum for each class of the library. Concerning distances, different types are computed: for example the Euclidean or the Mahalanobis ones [15].

In order to have supervised methods, a threshold needs to be defined. For example, if the correlation coefficient is higher than a certain threshold, the two compared spectra are considered as belonging to the same class. EMEA guidance encourages the application of wavelength correlation with a 95% threshold or the maximum wavelength distance [27]. However, other techniques are acceptable if justified.

2.1.2.2. Discriminant analysis. LDA is a linear and parametric method with discriminating characteristics [33]. LDA focuses on finding optimal boundaries between classes. LDA like PCA is a feature reduction method. However, while PCA selects a direction that retains maximal structure in a lower dimension among the data, LDA selects the directions that achieve a maximum separation among the different classes [34]. We can notice that LDA uses Euclidean distance to classify unknown samples. Quadratic discriminant analysis (QDA), a non-linear classification method [15] must be mentioned as this method proved its efficiency in solving particular NIR problems.

2.1.2.3. K nearest neighbours. K nearest neighbours [35,36] (KNN) is a non-parametric method. An unknown sample of

the validation set is classified according to the class to which belongs the majority of its K nearest neighbours in the training set [37]. The matrix of distances of the validation set samples to all other spectra of the training set is computed. The neighbours of an unknown sample are the samples having the lowest Euclidean norms. The predicted class is the class featuring the largest number of objects among the K neighbours.

The accuracy of the LDA, QDA and KNN method has been evaluated on a tablet data set and a capsule data set to classify samples of clinical studies [38]. In the same manner, Wu et al. compared several classification methods [39]. The selection of the most accurate classification method is data set dependent; therefore several chemometric methods have to be tried out.

2.1.2.4. Modelling methods. SIMCA [40] uses the modelling properties of principal component technique (PCA). It is a parametric method. This method considers each class separately. For each class, a PCA is performed which leads to a principal component (PC) model [10]. The validation set is used with every class models. An unknown sample is assigned to the class described by the model that produces the smallest residue during the prediction. SIMCA puts more emphasis on similarity within a class than on discrimination between classes. For example SIMCA has been applied to identify NIR spectra of 10 pharmaceutical excipients [41]. For each type of excipients at least 15 samples were collected in the data base, considering different batches and various suppliers.

PLS-DA [42,43] is a parametric and linear method. Partial least squares (PLS) identifies latent variables in the featured spaces which have maximal covariance within the predictor variables.

Some applications of this kind of methods were published: PLS-DA was applied to the supervised classification of samples out of dissolution specifications and the identity of blistered tablets was controlled on transmission spectra [44,45].

2.1.2.5. Non-linear methods. Artificial neural networks (ANN) are non-linear and non-parametric classification methods [46,47]. ANN are composed of several layers of neurons [22]: input, hidden and output layers. A neuron is a processing unit which transforms by an activation function input into an output data [48].

Several types of neural networks can be used for classification. Wang et al. [49] describe the advantages of multivariate discriminant analysis and feed-forward neural networks as classifiers. Two other networks are dedicated to the classifications: learning vector quantization neural network (LVQ) [50,51] and probabilistic neural network (PNN) [50].

Concerning pharmaceutical applications, networks [52] were applied to the identification of powder samples of sulfaguanidine based on diffuse reflectance spectra and their first derivative spectra. The networks were applied to discriminate qualified, unqualified and counterfeit powders.

Finally, during the last few years, a large number of applications deal with the support vector machines (SVM). In pharmaceutical application, SVM were applied for drug design [53,54]. However, no pharmaceutical applications of SVM with NIR spectra have been found.

2.2. Pharmaceutical applications

2.2.1. Identification and qualification

- **Analysis of starting materials**

International Conference on Harmonization (ICH) guidelines describe the importance of the identity tests [55–58]. Pharmacopoeias [59] have selected analytical methods to identify raw materials: for example HPLC, optical rotation, and colorimetry. The NIR application for qualitative analysis is also described by the European Pharmacopoeia in Chapter 2.2.40.

The identification of incoming raw materials is now a common NIRS application [60] thanks to the minimal sample preparation. A lot of publications describe this kind of applications because NIR can be applied to control excipients, active pharmaceutical ingredient (API) and final products. Ulmschneider et al. applied NIRS to identify different types of starches, sugars, celluloses, intermediates and active ingredients with principal component analysis and the cluster calibration module of the Nircal® software (Büchi AG) [61–64]. NIRS was applied to differentiate between the different Avicel products (microcrystalline celluloses PH-101, 102 and 200) [65]. The discrimination of celluloses [66] is statistically significant. Cellulose ethers were identified by NIR spectroscopy, however the separation of methylcellulose and cellulose ethers with methyl or hydroxyalkyl groups was not possible. Different types of polyvinylpyrrolidones (povidones) are characterized by their viscosity measured in water. Kreft et al. [67] have developed a NIRS method using SIMCA for the determination of the povidone types. The identification of raw material can be performed directly at-line at the reception in the warehouse or in the dispensing.

The NIR spectroscopy in comparison to wet chemical methods to control the quality of a pharmaceutical intermediate has been investigated. 7-Aminocephalosporanic acid has been chosen as an example by Andre [68].

A transflectance NIRS method [69] was developed for the identification of 15 solvents using a correlation coefficient. The optimum conditions for the clustering were obtained with the second derivative over the wavelength range 1136–2000 nm.

Strategies for development and optimization of libraries have been discussed by Gerhausser and Kovar [70]. Spectra of 17 benzodiazepines have been collected over the wavelength range from 1100 to 2500 nm. Validation of the models by predicting the identity of test set samples showed that the use of the correlation coefficient and the second derivative improved the recognition rate. Two pattern recognition methods, correlation coefficient and distance, were applied to confirm the identity of 117 drugs by using libraries based on full-range second derivative spectra. In this study, the construction of sub-libraries in order to discriminate similar drugs did not improve the classification.

After the identification, the qualification of raw materials needs to be presented. Qualification will determine whether a sample is within the normal variability range or is subject to over limit deviations. Distance based methods are often applied for the qualification of products. If the sample belongs to the same population as the reference product, then there is a probability of 99.7% that the distance will be less than three times the standard deviation. If the maximum distance is higher than that value, then the sample is from a different population.

- **Tablets**

  A method was developed for the identification of illegal ecstasy tablets [71]. The main identified substances were *N*-methyl-3,4-methylendioxyamphetamine, *N*-ethyl-3,4-methylendioxyamphetamine and amphetamine. The discrimination of the substances in tablet matrices was possible. Near infrared transmission spectroscopy combined with chemometric methods can also be applied for identity confirmation of clinical trial tablets [45]. Tablet identification is commonly performed by NIRS in quality control laboratories.

2.2.2. Polymorphism

The ability of a substance to occur in different crystalline forms is called polymorphism. The solid-state properties have an influence on the stability and dissolution properties of the pharmaceutical product. Moreover, fast analytical methods need to be implemented to help the galenical production. Infrared and Raman spectroscopies have been applied successfully to characterize API polymorphic forms. Nevertheless NIRS is also an analytical tool to solve this pharmaceutical issue.

The processing of NIR spectra provides information concerning the crystalline form of miokamycin [72]. NIR could increase the understanding of physical forms of theophylline [73]. The characterization and the analysis of azithromicin, an antibiotic derived from erythromycin A, was studied by Blanco et al. [74]. In this study, the NIRS method was validated according to the ICH guideline.

The suitability of near infrared spectroscopy to follow changes in both the amorphous and crystalline lactose at room temperature has been investigated [75] and their differentiation found possible by studying the NIR frequencies of water peaks.

2.2.3. Other applications

After the presentation of the main applications of NIRS, other studies can be mentioned as well. Yoon et al. [76] discriminated the production sites of proprietary tablets. The PCA score plots showed that spectra of tablets originating from different manufacturing sites are statistically different. The identification of production sites is valuable to manufacturers, customers and industry regulations. Scafi and Pasquini [77] used NIRS to identify counterfeit drugs. The identification is based on the comparison of the NIR spectrum of a sample with spectra of the authentic drug using PCA and SIMCA methods. The results showed that at least 50 spectra must be included in the training set in order to obtain 100% of good classification. Vredenbregt et al. used NIRS to determine authenticity and origin of Viagra® tablets [78]. A new development in pharmaceutical quality control is the use of a combination of analytical methods to make a fingerprint of a drug substance or mixture [79,80]. The food and drug administration (FDA) is paying attention to the NIRS as analytic tool to fight counterfeiting and to detect non-conformity with the original product.

NIRS and chemometrics were also performed to understand process and dissolution issues [81]. More precisely, then it showed how NIRS and IR imaging can be useful in understanding batch differences due to different process conditions. Beginning with a qualitative analysis of the potential applications of NIRS and IR imaging to solid forms, the ability of NIRS to detect the effects of melt granulation time-temperature gradient, compaction force, coating formulation and coating time were tested on pilot production samples.

One important operation in manufacturing solid pharmaceuticals is the monitoring of the homogenization process. An uniform distribution of the active ingredient and the excipients in a pharmaceutical blend is essential to have the correct dosage [82]. Details will be given in the on-line part of this review. As a new field of application, the NIR qualitative analyses for biotechnology products is also raising up. For example NIRS was applied to identify bacteria stains [83,84]. However, most of the applications of NIRS in biotechnology field are quantitative and will be presented in the next paragraphs.

2.3. Practical aspects to develop qualitative methods with NIRS

The classification methods and the main qualitative applications were presented. Thus, the aim of this part is to provide practical aspects for developing a NIR qualitative method, optimizing calibration, managing and validating a library.

- **Sampling and data pre-processing**

  Yoon et al. [85,86] optimize the sample presentation of pharmaceutical excipients for the NIR measurements. Using a Foss NIR Systems Rapid Content Analyzer, three parameters were identified with significant impact on the classification algorithm: cup diameter, sample thickness, and cup material.

  The NIR spectra can be affected by the particle sizes of the sample, by variations of the optical pathlength and by crystalline forms. That is why a well defined sample analysis preparation protocol is required. To avoid or decrease these interferences mathematical pretreatments are applied on the spectra.

  The most current data pretreatments are the normalization methods like standard normal variate [87] (SNV) and multiplicative scatter correction [88] (MSC), the derivative methods (for example the Savitzky-Golay method) and the Orthogonal Signal Correction (OSC).

  The spectral pretreatments and the selection of the wavelength ranges should be carefully chosen before applying the pattern recognition methods in order to optimize the model. de Groot et al. [89] and Wu et al. [90] have shown the influence of wavelength selection and data pre-processing on NIR classification. The effect of data pre-processing (no pre-processing, offset correction, de-trending, SNV, SNV+de-trending, MSC,

first and second derivative) on the identification of 10 pharmaceutical excipients was also investigated by Candolfi et al. [91].

- **Library management (calibration set management)**

  The identification process involves two steps, *viz.* recording NIR spectra for the calibration set, a so-called 'spectral library', and then for the validation set. Correct identification relies on the choice of the calibration set, i.e., spectra to be included in the library. The spectra should contain every possible source of variability associated with the product and the manufacturing process (i.e., different batches and different storage times). Spectral variability can also be considered by including spectra of the same sample recorded by different operators on different days.

  It is difficult to determine the number of spectra to be included in a library. The number of samples will depend on the study. For a product manufactured in a reproducible way, variability can be taken into account by samples from 5 to 10 different batches. However, if manufacturing reproducibility is poor, the number of required spectra to be included can dramatically increase. The EMEA guideline [27] advises to use at least five batches: three for the calibration and two for the validation to identify excipients by NIRS.

  Blanco and Romero [92] has described the theoretical and practical aspects of library construction. The procedure is demonstrated by constructing a library including NIR spectra of 125 different raw materials using the correlation coefficient as classification method.

- **Calibration transfer and model update**

  After model development, calibration needs to be maintained. Possible new variability sources have to be included in the model. That is why models are regularly updated and the calibration quality checked. The effect of model updating [93] on the identification of a pharmaceutical excipients based on its NIR spectra has been investigated. An updating approach consisting in adding newly available samples to the training set and rebuilding the classification model was applied.

  Calibration is time consuming. Thus the same model is often applied on several spectrometers in order to save time. Ulmschneider et al. [61] have successfully built transferable cluster calibrations for the identification of different solid excipients with near infrared spectroscopy. A review discusses some of the causes of the non-transferability of NIR spectra for the identification of pharmaceutical excipients [94]. Calibration transfer method will be described in the online part of the review. However, we can mention that the qualitative calibration transfer appears to be easier than the quantitative calibration transfer.

- **Pharmaceutical validation**

  A practical experience with the EMEA Guidance on the use of near infrared spectroscopy by the pharmaceutical industry has been published [95]. The practicability of this guidance for the identity test of starting materials in a generic manufacturing site was explored.

  Without chemometrics NIRS would not be an analytical method. It is important when a NIR device is purchased to check the chemometric tools included in the software package and it is necessary to evaluate the potential of the new chemometric methods for the pharmaceutical application. The software must also be CFR 21 part 11 compliant to be implemented in a GMP environment.

3. Quantitative analyses by near infrared spectroscopy

Once the classification of samples has been achieved it can be useful to know more precisely in what extend samples are different. Therefore, the development of quantitative model appears necessary. Historically the first quantitative determinations were performed on the moisture of samples, thanks to the two strong water bands absorbing at 1450 and 1940 nm [2]. The quantitative part of this review will be subdivided into three parts. A first one deals with the regression methods, i.e., the chemometric tools used to construct the model, then a second part will present a review of the application in the pharmaceutical environment and a final part will expose the practical aspect to construct a quantitative analysis by NIRS, i.e., the selection of the samples and of the statistical indicators.

3.1. Regression methods

First of all the Beer-Lambert law is of course the easiest way of constructing a regression line. This law is considered as well known and due to its difficulty to apply it to NIR spectroscopy it will not be discussed here.

3.1.1. Multi-linear regression

The Multi-linear regression (MLR) [16] is the oldest of the presented methods and is less and less used in applications due to the improvement of computation power. This regression allows establishing a link between a reduced number of wavelengths (or wavenumber) and a property of the samples. The prediction y_j of the search property can then be described with the formula:

$$y_j = b_0 + \sum_{i=1}^{k} b_i x_i + e_{i,j}$$

where b_i is the computed coefficient, x_i the absorbances at each considered wavelength and $e_{i,j}$ is the error.

Each wavelength is studied one after the other and correlated with the studied property. The selection is based on the predictive ability of the wavelength. The three modes of selection are: forward, backward, and stepwise. When the correlation reaches a value fixed by the operator it is kept as a part of the model calibration wavelengths. The model is then computed between this set of calibration wavelengths and the reference values of the studied property.

3.1.2. Principal component regression

The principal component regression (PCR) is divided into two steps. First the spectral data are treated with a PCA. Then a MLR is performed on the scores as predictive variables [96].

The prediction equation is written $\mathbf{Y}_{\text{sampling}} = \mathbf{T}_{\text{sampling}}\mathbf{b}$ with $\mathbf{T}$ is the new dimensional coordinates, $\mathbf{Y}_{\text{sampling}}$ the reference values and $\mathbf{b}$ is the coefficient vector.

There are several advantages in the use of this method. The PCA suppresses the spectral colinearity. Meanwhile there is no guarantee that the computed principal components are correlated to the studied property.

3.1.3. Partial least squares regression

In PLS [97] method the regressions are computed with least squares algorithms. The goal of the PLS is to establish a linear link between two matrices, the spectral data **X** and the reference values **Y**. This technique is modelling both **X** and **Y** in order to find out the variables in **X** matrix that will best describe the **Y** matrix. This can be explained by the representation of the spectra in the space of wavelengths in order to show directions that will be linear combinations of wavelengths called factors which describe best the studied property.

PLS had enhancements during the years: for examples they became O-PLS [98] when combined with OSC corrections [99,100]. Another enhancement is the moving window PLS [101].

PLS has the PCR advantages without the drawbacks thanks to the latent variable selection according to the covariance matrix between the data and the investigated parameter.

3.1.4. Artificial neural networks

ANN are made of several layers. The first one, called the inputs, is the entry variables which are absorbances at specific wavelengths. The outputs which are in our case a content value also represent a layer of artificial neurons. The so called "hidden layers" represent the modelling process. They allow getting the output values from the first layer of neurons. To compute this, hidden layers need to be trained by back-propagation, i.e., computing the transfer functions from the output to the entries to achieve the necessary prediction skills of the network. Once the neural network has been trained it is used on the entry variables to get the predictions.

The use of ANN requires the optimization of a lot of parameters such as the number of hidden neurons or the number of iteration to train the network, and of course the selection of the data pretreatment and the selection of the wavelengths [102,103].

This calibration method is used by Plumb et al. to determine the effect of experiment design on the modelling of tablet coating [104] and in 2005 he has compared several training algorithms to predict dissolution rate [105]. The major advantage of the ANN is their ability to construct models around non-linear relationship between the measured data and the predicted properties.

3.1.5. Support vector machines

As classical least squares methods the SVM or support vector regression (SVR) finds a linear relation between the regressors and the dependant variables [106]. The cost parameter to determine the best model is different from the one used in the former regression methods, that is why SVR can be applied to non-linear phenomena [107,108]. These two new techniques have been compared in the study of Thissen et al. [106]. Blanco et al. have compared several of these regression method (mainly PLS and ANN) in an article [109]. As for the qualitative application, the selection of a quantitative method is dependant of the study and no general rules for the selection of method can be provided.

This part as well as the previous one showed that the use of chemometrics is a solution for performing either qualitative either quantitative analyses. Chemometrics are anyway a dangerous tool and one should always bear in mind what he is doing, what he is searching for. It is a tool that has to be used with parsimony, a recent issue of Trends in Analytical Chemistry [110] gives a good view of "*use and abuse of chemometrics*".

3.2. Pharmaceutical applications

3.2.1. Physical parameters

NIR spectra contain information about the chemistry and the physical properties of samples. Consequently, various pharmaceutical parameters can be quantitatively analyzed by NIRS such as hardness, particle size, compaction force, and dissolution rate [111].

NIRS is now used in a pharmaceutical environment to determine a large panel of physical parameters on powders as well as on tablets. Hardness of tablets is determined in different studies with the well established regression methods (PLS, MLR). Morisseau and Rhodes used MLR and PLS to predict the hardness of tablets. The accuracy of the results are highly dependent on the products and their formulations [112]. PCR and the slope of the best-fit line, thereby the spectrum is reduced to a slope and an intercept, are described in a study published in 1999 [113]. Standard error of prediction (SEP) of 0.477 and 0.539% are achieved on two different models computed by artificial neural network in a study on drug content and tablet hardness by Chen [103]. The correlation of compression force and NIR spectra at a specific wavelength is shown by Guo et al. [114]. Blanco and Alcala have more recently shown the possibility to predict the pressure of compaction on a laboratory sample by using a PLS model [115].

The use of different regression methods (linear, quadratic, cubic and PLS) allows following the percentage of the drug released in the medium by a tablet. The dissolution profile is determined by NIRS on six different times between 15 and 120 min [116].

Berntsson has produced results on the determination of effective sample size when analysing powder blends with NIR reflectance spectroscopy, showing that on powder the analyzed depth is not further than 0.75 mm and depends of particle size [117,118]. Otsuka published in 2004 results on the scattering effect due to particle size measured by using a PCR model [119].

3.2.2. Polymorphs determination

The polymorphic form of a product is a key parameter of this product as it can modify the dissolution properties of the final drug. But it also allows the detection of possible counterfeits, therefore insuring the correct polymorphic form of a proprietary product appears necessary.

The determinations of ratio between amorphous and crystalline forms of products are usually done by X-ray diffraction. Several studies showed that this measurement can also be done by NIR spectroscopy and that the limit of detection of

amorphous forms in crystalline form is lower by X-ray than by NIR [120–122]. Bay observed great agreement between NIRS and X-ray on determination of glycine crystallinity. The standard error of prediction (SEP) of NIRS is 3.2% and it detects crystallized glycine at a lower rate than X-ray diffraction [123]. The NIRS combined with different regression methods is used in several polymorphism or crystallization applications on numerous products for example by Berntsson, Févotte, or Patel [72,74,124–127]. The NIRS can also be used as described by Savolainen et al. for evaluation of the differences in the amorphous state of indomethacin depending on the starting crystalline form and the preparation process [128].

3.2.3. Moisture determination

The moisture determination is one of the very first applications of NIRS in the pharmaceutical environment. Water is a critical parameter that has to be ensured in a lot of pharmaceuticals because it is a key compound for the stability of the product. The study of this compound is mainly due to the importance of the water signal in the NIR spectral range thanks to two different bands centered at 1450 nm and around 1940 nm. NIR spectroscopy is used to determine the water content in powders or granulates [129–131], tablets or capsules [132,133], as well as in lyophilised vials or in solutions [134]. As it is being developed for a long time now, most of the relevant applications of NIRS in moisture determination are on-line, and will be described in the corresponding part of the review.

3.2.4. Content determination

A lot of studies have been published during the last few years concerning determination of chemical compound content such as API, excipients or moisture in pharmaceuticals. Meanwhile many of the studies are shortly presented here.

Samples can be of various types, e.g., be powders, granulates, tablets, liquids, gels, films or lyophilised vials. One original study has shown the determination of ethanol, propylene glycol and water directly through amber plastic bottles [134].

NIR spectroscopy is used for the quantitative determination of active, excipients, moisture or coating thickness (Table 1).

A study compares NIR spectrometers, FT-NIR, FTIR-PAS (Fourier Transform Infrared-Photoacoustic Spectroscopy), FTIR-ATR (Fourier Transform Infrared Attenuated Total Reflectance), Diffuse Reflectance Infrared Fourier Transform Spectroscopy and FT-Raman applied for the determination of vitamin C in powders and solutions [135].

Few studies are comparing the transmittance and reflectance technology in NIR spectroscopy for quantitative determinations [136–139].

A more recent study of Chalus et al. compares different data pretreatments and regressions to compute models for prediction of active in low-dosage tablets [140], another study compares different NIR spectrometers for the determination of active content in tablets [141].

More and more studies are published on the use of NIR spectroscopy to follow the process of production of tablets, from the raw materials to the end product, either coated or not and even on packaged tablets [81,142,143].

Feng and Hu have shown in a recent study that building a single universal calibration for determination of API in different manufacturers' tablets was achievable. The study showed the results on two different active ingredients [169].

As shown by the table the quantitative applications are numerous. Today a new field of pharmaceutical application emerges: the products obtained via biotechnology also represent a domain of expansion for NIRS.

3.2.5. Lyophilisation

Lyophilisation is a widely used method for the formulation of a wide range of pharmaceutical products particularly susceptible to degradation in aqueous solutions like peptides, proteins or complex organic molecules. The aim of lyophilisation is to produce substances with good shelf stability and which are unchanged after reconstitution with water. The first NIRS application depicted for lyophilised products was the determination of residual moisture content through intact glass vials [170–173]. The advantage of this technique in comparison to the traditional methods like Karl Fischer titration (KF), thermogravimetry (TG) or gas chromatography (GC) is that it is rapid, non-invasive and non-destructive. It avoids opening the vials and risking a contamination from the atmospheric moisture which can result in error in the determination of residual water content. NIR represents thus an alternative approach for the quality control of lyophilised pharmaceuticals. Nevertheless classical methods such as KF are required as reference methods for setting up the different NIR calibrations.

First Kamat and DeLuca [171] have described the suitability of NIR for the determination of residual moisture content in lyophilised sucrose. Last and Prebble [170] have made an assessment of the applicability of the moisture calibration for the humidity prediction in a product containing different amounts of active ingredient. Jones et al. [172] have done it for the transferability of this kind of calibrations between instruments and sites. Savage et al. [173] have compared two techniques to determine moisture content during dry-heat viral inactivation; the gravimetric method Loss on Drying (LOD) and the KF titration versus NIRS.

Recently Lin and Hsu [174] have examined the application of NIR not only for the determination of the residual moisture content, but they additionally have investigated how changes in product configuration like cake porosity, cake dimensions and excipient-to-protein-ratio could affect the accuracy of NIRS residual moisture content prediction.

A lot of methods focus normally on the total amount of residual moisture, but in a recent study Cao et al. [175] furthermore have found out that by means of NIRS it is also possible to differentiate and quantify water of different energetic states. In this work quantitative methods, based on curve fitting analysis and PLS regression models have been developed to quantify both hydrate and surface water content in lyophilised mannitol. Knowledge about hydrate water in lyophilised pharmaceuticals is very important; as it is released during the storage it can have a significant impact on the stability of the formulation. The information about the combination of both surface and hydrate water is also required. As shown by Cao the hydrate peaks in the

Table 1
Quantitative determination by NIRS

Analyte	Sample form	Regression method	Remarks	Ref.
Oxytetracycline	Oxytetracycline base	PLS, PCR	The PLS is preferred to PCR for the calibration of assay. The assay is ranging from 91.95 to 94.43%. The SEP is 0.285%	[103]
St. John's wort	Plant	PLS	Hyperforin and biapigenin are calibrated in this study. Dry extract of St. John's wort are measured in reflectance. For the hyperforin calibration the RMSEP is 0.22% for a content range of 1.0–6.0% and for biapigenin the RMSEP is 0.024% for a content range of 0.20–0.55%	[144]
Ketoprofen	Translucent gel	MLR, PLS	MLR is preferred to PLS because of its simplicity and because it gives a good clue on the ability of NIR to predict content for these formulations	[145]
Nystatin metronidazole	Cream	PLS	The two different ingredients of the cream are determined by means of PLS models. The measurements are made through the bottom of the vials. Both models have a correlation coefficient greater than 0.998	[146]
Testosterone	Thin films	PCR	Two layers mucoadhesive thin film composites disks for release of testosterone are measure by NIRS. The reached SEP is 0.17 mg for a content ranging from 0 to 4 mg	[147]
Selamectin moisture	Topical formulation	PLS	Two models have been computed. One for the simultaneous determination of selamectin and moisture and one specially designed for the determination of moisture	[148]
Citral	Lemon grass, lemon oils	MLR	The citral content in lemongrass and lemon oils are ranging between 70 and 77% and between 2 and 16%, respectively. The SEP obtained for MLR calibration are 0.48% and 0.12% in lemongrass and lemon oils	[149]
Cineol	Eucalyptus oil	MLR	The eucalyptol content is ranging from 75 to 85% (w/w). The best MLR calibration reached a mean accuracy of 0.86% on the validation set	[150]
Potency lipids	Monensin fermentation broth	MLR	NIRS appears to be more precise than the laboratory reference method for both products. Nevertheless the study is limited by a lack of stability	[151]
Amilose	Starch	Peak ratio	The amylose was ranging from 2 to 95%. The computed model presents a RMSEP of 1.2%	[152]
Clotrimazole	Extruded film	PLS	The drug content is ranging from 0 to 20% in a hot melt extruded film of polyethylene oxide. (RMSECV = 0.298)	[153]
Water	Lyophilised samples	PLS	Two PLS models are built to follow the drying process. The first one is built for water content between 1 and 40% (w/w) with a SEP of 1.85% while the second one is constructed on samples ranging from 1 to 10% (w/w) water with a SEP of 0.42%	[154]
Kavapyrones Kavain water	Extracts of piper methysticum forst	PLS	For all the calibration the SEPs are lower than 0.094%	[155]
Coating thickness	Tablets	PLS	The measurements are made directly in a fluidized bed. The thickness of the coating is followed with a correlation coefficient of 0.97 and a RMSEC of 2.2 μm for a thickness varying between 0 and 50 μm	[156]
Coating thickness	Tablets	PLS	The tablets are composed of two halves of different chemical composition. Models are computed for each half and for both. In the first case performing a PCA is necessary to determine which half have been measured	[157]
Metformin	Tablets	PLS Linear regression at 1 λ	PLS appeared to be more accurate than single wavelength regression. The standard error is ±1.56%	[158]
Caffeine	Tablets	PLS	The range of caffeine content is 0–100% (m/m) to construct the calibration for analysing tablet of nominal content 58.82% (m/m). The SEP on linearity validation is 1.37% (m/m)	[159]
Steroid	Tablets	PLS	The tablets cover a range from 2.94 to 17.64% (m/m) of active. The tablets are measured in transmittance, the SEP allowed the use of NIRS for the assay of tablets for batch release	[160]
Gemfibrozil	Tablets	PLS	Tablet of two related pharmaceutical preparations are first identified by a classification model and then predicted with a sole calibration. This calibration allows SEPs of 1.3–1.6% for the cores and coated tablets of the formulation with 751 mg/g of active and SEPs of 0.7–1.2% for the cores and coated tablets of the formulation with a content of 810 mg/g	[161]

Table 1 (*Continued*)

Analyte	Sample form	Regression method	Remarks	Ref.
Ibuprofen 800 mg	Thick tablets (1150 mg)	PLS	The tablets of 7.6 mm thickness have to be re-pressed to be reduced to 3.6 mm thick. The measurement in transmittance is thus usable on an especially dedicated device. The drug content ranged from 752 to 848 mg	[162]
Paracetamol	Tablets	MLR, PLS	Two wavelength selection modes were tried for the MLR. The computed models present SEP between 0.6 and 1.0% depending on the pretreatment. The PLSR model reached SEP of 0.6 for most of the pretreatment	[163]
Paracetamol	Tablets	MLR	The MLR model is computed on two wavelengths. The SEP for this model is 0.71% (m/m)	[164]
Acetylsalicylic acid	Effervescent tablets	PLS	This study assays acetylsalicylic acid in three different formulations were it can be either the only active or combined with vitamin C or with vitamin C and paracetamol. The measurements are performed in reflectance and in transmittance on intact tablets and in reflectance on milled tablets	[137]
Paracetamol amantadine hydrochloride	Tablets powder	ANN	The assays of paracetamol and amantadine hydroxide are simultaneously determined by an ANN model. Models have been built with different pretreatments and on tablets as on powders	[102]
Paracetamol diphenhydramine hydrochloride caffeine	Powder	ANN, PLS	The study compares different pretreatments and PLS to ANN. ANN improves the results compared to classical PLS	[165]
Ferrous lactate dihydrate	Granulate-Powder	PLS	The concentration range was 650–850 mg/g. Identification is first performed on the samples. The laboratory samples are powders while production ones are granulates. Both types are included in the calibration. The reached SEP is 3.8%	[166]
Diphenhydramine	Tablet transmittance powder reflectance	PLS	Tablets of diphenhydramine are measured in reflectance and in transmittance. Their milled form is measured in reflectance. Results are comparable for the three kind of measurements even if results in reflectance on intact tablets are lower	[167]
Dexeketoprofen	Hydrogel	PLS	The concentration range was first from 9 to 15 mg/g, then placebos are added to the model. The relative error for prediction is 1.6%	[168]
Mirtazapine	Powder tablet	PLS	The PLS model for determination content is based on laboratory powder samples and production tablets. The first factors of the model had to be excluded as they were explaining the physical differences of the samples. The final selected model used four PLS factors	[115]

spectra decreased with the time, demonstrating that the hydrate form of mannitol is unstable, and as a result of the dehydration subsequently the surface water increased in the lyophilised cake.

Regarding the stability aspect of lyophilisates, studies of NIRS applicability have been carried out by Stockvold et al. [176]. A clear relationship between moisture, storage temperature and time has been found. Moisture content in freeze-dried product increased with storage time and temperature due to the water released from the stoppers. Derksen et al. [177] have found, based on the relation between the residual moisture content and the content of the active ingredient, combined with the Arrhenius relationship between degradation rate constant and temperature, that it was possible to predict the moisture content specification for product shelf life.

The suitability of NIR as alternative to X-ray powder diffractometry (XRD), differential scanning calorimetry (DSC), freeze-drying microscopy, nuclear magnetic resonance (NMR), Fourier transform infrared (FT-IR) and Raman spectroscopies to determine the degree of crystallinity of pharmaceuticals has been proven by Bai et al. [123]. The mentioned methods used in the routine present significant drawbacks, such as sample size, sample preparation and measurement times. The authors have demonstrated that NIRS compared to the reference method XRD is well suited for quantify glycine crystallinity in lyophilised cakes. For providing structural support to the final lyophilised product, glycine is desired to be in a highly crystalline state.

NIRS has also been found to be a good tool for evaluating protein structure. It presents several advantages in relation to other standard tools for characterization of lyophilised protein formulation as, e.g., FT-IR. As no sample preparation is required, the possible vapor absorption or structural changes in the protein due to the KBr pellet preparation can be avoided. Bai et al. [178] have shown for proteins in the solid-state that both drying-induced damage and thermally induced denaturing can be detected by means of NIRS. It has also been possible to assign well defined absorption bands in the NIR spectrum to α-helix and β-sheet. Proteins can be distinguished based on well defined bands in the NIR spectrum. Izutsu et al. [179] also have investigated the secondary structure of seven different proteins in various physical states. The NIR spectra of the proteins in aque-

ous solutions and in freeze-dried solids have shown that NIR not only enables the detection of structural differences between proteins with identical amino acid compositions but also the detection of higher-order structure of the proteins.

In another study Izutsu et al. [180] have elucidated the effect of counter ions on the physical properties of L-arginine in various physical states by using NIR. While some acids and salts raised the glass transition temperature (T'_g) of the L-arginine, which is desirable to avoid cake collapse during the drying process and to guaranty a good stability of the amorphous solid after drying, others had little effect or lowered it. The NIR spectra analyzed have revealed the presence of interactions around the L-arginine amino or guanidyl groups with counter ions and in addition have proven that NIRS is also a useful technique for the detection of changes in the physicochemical properties of amorphous freeze-dried solids.

In summary all these applications have shown that NIRS has a big potential for the quality control in development and production of freeze-dried products [181,182]. Not only due to the specific information taken out of each application but also due to the additional physicochemical information carried out with the spectra. A new dimension on the application of NIR to freeze-dried products has been provided by Brülls et al. [183]. A real time in situ monitoring of the lyophilisation process has been performed using a NIR probe and a fine wire thermocouple both placed in the center of a 10 ml vial, 1 mm above the bottom. The real time monitoring has provided important new information about physical changes during the lyophilisation process, such as freezing (the start of ice formation, completeness of ice formation) and sublimation (transition from frozen solution to material free from ice) in agreement with the data from the thermocouple. But also information which was not possible to detect with classical [173] process monitoring techniques such as the rate of desorption and the end point of the drying process has been provided.

3.2.6. Biotechnology

Biotechnology products are raising more and more in the pharmaceutical industry. Therefore, analytical methods have to adapt to these new features. The NIRS is one of the most flexible technologies and this is why a lot of studies are considering the analysis of bioprocesses and their productivity. In 2000, Arnold et al. tried to monitor at-line the fermentation of *Streptomyces fradiae* to produce antibiotic tylosin. This process involves the use of two carbon sources (glucose and methyl oleate), glutamate, and ammonium as nitrogen sources. Spectra were acquired in the transmittance mode. Models were built for this four analytes using PLS and the second derivatives spectra. Then models have been validated and provide accurate results [184]. In the same type of fermentation, Vaidyanathan et al. have compared the models developed using two measurement modes: diffuse reflectance and diffuse transmittance. Both measurement modes have been employed in the bioprocess. Models were developed by PLS for oil and tylosin. Normally, diffuse reflectance geometry offers less information than transmittance measurements. The spectral information in reflectance mode appeared restricted. This is important for further on-line application [185]. In another article Vaidyanathan et al. have investigated a bioprocess of *Penicillium chrysogenum* which produces penicillin G. NIR spectra were acquired at-line in transmittance mode for biomass, total sugars, ammonium, penicillin and extracellular protein. Models were computed with a PLS and assessed with SEP. Despite the filamentous nature of the biomass, multivariate models could be developed [186]. The robustness of models was then validated by challenging them with analyte and matrix variations. Models are affected by changes in the matrix or analyte concentration. The use of an external validation set facilitated the identification of the relevant models [187]. Frazier et al. have characterized the effective diffusivity of glutamine in immobilized cell materials. Glutamine was selected as a critical metabolite for animal cell cultivations. Spectra were acquired in transmittance for each of the 10 positions of pure agarose solutions in the chamber. This was repeated with 11 gel samples. PLS-model was computed and evaluated with SEP. Diffusivity was calculated for the entire gel [188]. The chemical environment of mammalian cell cultures must be controlled to maximize productivity. Rhiel et al. used human prostate cancer cells. Glucose and glutamine are consumed whereas lactate and ammonia are produced in a serum-based medium. Spectra were collected in transmittance. Models were built for analytes with a PLS. These results are to be introduced for on-line monitoring [189]. In a fibroblast culture, determination of glucose, lactic acid and ammonia were considered. Spectra were measured for aqueous mixtures and media samples. PLS regression allowed constructing accurate models [190]. Such works enhance the at-line application of NIRS and encourage further on-line application.

3.2.7. Other pharmaceutical applications

Some quantitative studies are not directly on the pharmaceutical product nevertheless they can be linked to pharmaceutical industry. Laasonen has shown the possibility to determine the thickness of the plastic packaging of pharmaceuticals. The SEP for this determination was only 4.3 μm [191]. Such works enhance the at-line application of NIRS and encourage further on-line application.

3.3. Practical aspect: sample selection for calibration and validation sets

NIR spectroscopy cannot be considered as a primary method. Therefore, to develop a quantitative model it is necessary to have a reference method to evaluate the property of samples which will be used in the computation. These wet chemical methods give what is called the reference values for the samples. These values will be used to compute the model.

The aim of computing quantitative models is predicting a property of unknown samples with their near infrared spectra. A model is built and validated by using several sample sets. A first one is the calibration set used to compute the model. A second sample set is the validation set used to evaluate the ability of the model to predict unknown samples (Fig. 1). The calibration and the validation set have to be independent, they must consist of samples from different batches.

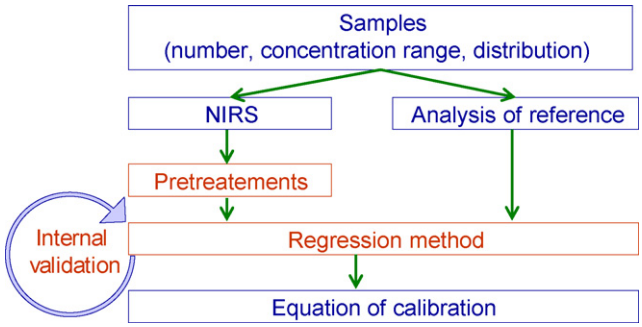

Fig. 1. Scheme for the construction of a quantitative model.

Once the model has been constructed and validated it can be run as routine. It can be run on the device which was used during the development step or on another device. In this case one has to ensure the transferability of the model from one device to another one. Most of the time adjustments are necessary. The implementation of laboratory work to on-line is one of the main issues where transferability has to be faced. Calibration transfer methods and issues will be developed in the last part of this paper.

Before computing quantitative models, one has to think of what will be the aim of the calibration and what will be its minimal accuracy and limits of validity. Therefore, it is necessary to design an adequate range of samples comprising enough side variations to allow the future model to be robust [192–194]. The most common repartition will be that two third of the samples are used in the calibration set and one third is used as the internal validation set. Some studies on sample selection are comparing the use Kennard-stone, successive projections algorithm, random sampling and full cross-validation on modelling with MLR and PLS [195,196]. Wu et al. made a study on the influence of sample selection in the sets on neural networks models [197]. The full cross-validation, the leave on out method, can also be used for intern validation meanwhile this method is preferred when the number of samples is reduced.

Goodness of fit of a prediction model can be evaluated according to the following criteria: low Standard Error of Calibration (SEC), low SEP, high correlation coefficient (R^2), and low bias. SEC, SECV (standard error of cross-validation), SEP, bias, slope, and SEP(C) (SEP with bias correction) are criteria to evaluate the accuracy of the model. The formulae and statistical strategy are described by Naes et al. [198].

4. On-line control by means of near infrared spectroscopy

4.1. Pharmaceutical applications

4.1.1. Powder blending

The blending of API with excipients is a critical step in the manufacturing of pharmaceutical solid dosage forms. Without a homogenous blend it is impossible to get uniform doses with the right content of API later in the production process. However, the determination of blend homogeneity is problematic. Mostly at the moment, samples are removed from the blender bin by using a sample thief and then analyzed by conventional chemical methods, such as HPLC or UV–vis-spectroscopy. Apart from the fact that only the distribution of API is determined and the homogenous distribution of the excipients is assumed if the API is homogenously distributed, the sampling by a thief often changes the distribution of the powders and is thus associated with significant sampling errors. Moreover, the classical chemical methods are destructive, time and cost consuming—they are labour intensive, need solvents and are responsible for long cycle times as they are performed off-line. Therefore, the use of a fast, non-destructive method to determine blend homogeneity is favourable. NIRS offers those advantages, furthermore it observes all contents of the powder mixture, not only the API. As it is fast enough for real time analysis and non-invasive, an on-line or in-line application is possible, not only for determination of homogeneity but also for end point determination. A lot of studies have been carried out to explore the use of NIRS for powder blending process control. That NIRS has a great potential in powder blend analysis in principle has been shown by Wargo and Drennen [199–202]. Cho et al. have dealt with the effective mass that is sampled by NIR fibre-optic reflectance probes in blending processes and have shown that the sampled mass meets the requirements of FDA regulations [203]. Hailey et al. have demonstrated by using a NIRS fiber-optic reflectance probe in either a y-cone- or a bin-blender in combination with a graphical user interface and appropriate software that it is possible to use NIRS for in-line blend analysis [204]. Sekulic et al. have also evaluated the use of NIRS for on-line monitoring of powder blending processes by using a NIRS fibre-optic reflectance probe and showed its feasibility with a model-free approach [205]. Not only NIRS but also NIR imaging has been used by El-Hagrasy et al. who have also demonstrated the possibility to use NIR for on-line blending control, however they have pointed out that multiple sampling points are necessary for correct process control [206]. Sekulic et al. have focused on qualitative approaches to blend evaluation in a study using a small Flobin blender and a reflectance fibre optic probe. Different blends have been produced, monitored via NIRS and on the resulting data different mathematical pre-processings have been performed [207]. Skibsted et al. have presented a qualitative and a quantitative method. They have developed control charts and have thus been able to monitor the homogeneity of the blend [208]. In a more recent study they have gone further on by showing the use of NIRS applications on a complete manufacturing process in order to allow real-time release of products [209]. NIRS to quantify the drug content in a blend has also been used by Popo et al., however they have not taken spectra of the blend directly in the blender but have used samples they obtained by stream-sampling [210]. Berntsson et al. have described the quantitative in-line monitoring of powder blending in a Nauta mixer, both at laboratory and production scale. By high speed sampling, both average content and distribution of the mixture content have been assessed [211].

4.1.2. Granulation

In many cases in the manufacturing of solids, powder blending is followed by a granulation step. Granulates are often

necessary for later tablet compression or capsules filling, or they are the final drug form itself. They are produced by either dry granulation such as roller compaction or wet granulation such as fluid bed spray granulation or high shear mixer granulation. A central parameter is the moisture content of the granulate during granulation as it describes the granule growth kinetics. It is also important for the later process as it influences the properties of the granulate and thus, e.g., the hardening of tablets during storage. Classical measurement methods such as infrared dryers for moisture content determination need time and do not only slow down the manufacturing by waiting periods but are also too slow for process control. As NIRS responds in real time, the process might be monitored more efficiently, resulting in greater process reliability and optimized product characteristics. Rantanen et al. have used NIR-reflectance spectroscopy for in-line moisture content determination in fluidized bed granulation. They have monitored spraying and drying phases and have been able to determine drying end points [131,212]. The effects of binder and particle size on moisture determination by NIRS have also been investigated [213]. A non-linear calibration model has been developed by combination of NIRS with other process measurements [214]. Frake et al. have also applied in-line NIRS to a fluidized bed granulation to control the granule moisture content and changes in particle size [215]. Findlay and et al. have shown that by NIRS it is possible to control a fluidized bed granulation, they have determined not only the drying end point but also the time point when binder addition should be stopped [129]. Gupta et al. have investigated the use of real-time NIRS for process control of roller compaction. They have shown the link between the best fit line through the NIR spectrum and the strength of the compact and also determined the particle size distribution after milling of the compacts [216]. In another study they have determined content uniformity, moisture content and strength of compacts [217].

4.1.3. Drying

Drying is most of the time a critical step in the manufacture of pharmaceutical products. It is used in processes such as granulation (commented in the corresponding part of the review) or lyophilisation.

The current pharmacopeias establish a series of quality parameters for the release of lyophilised products, amongst others the destructive and time consuming KF titration. Only a small spot check is required to prove the compliance of quality. This few samples are representative for the whole batch and decisive for the release. Even when these samples are found to be conforming to the requirements, this does not guarantee the fulfillment for the rest of the products not controlled. Thus, an uncertainty of compliance is surrounding these products. A new trend is the use of in-line methods for a complete batch inspection for guarantee a homogeneous quality of the product avoiding the costly loss of complete batches in the case that some vials out of specification are found. Sukowski and Ulmschneider [182] have opened a new way to accurately predict in-line the residual moisture content in lyophilised products. In his feasibility study Sukowski has proven that with applying classical chemometrics (PCA for the identification of the product and PLS for the quantification) a 100% control of an entire batch at a line speed of 300 vials per minute is possible. As already mentioned in the lyophilisation part of this review, Brülls et al. have present a very promising way to survey qualitatively and quantitatively the drying of samples during a lyophilisation process [183].

4.1.4. Cristallinity

During the drying phase of wet granulation polymorphic changes can occur in the active ingredient or in some excipients. The polymorphic changes of glycine involve important changes in the hydrogen bonding of the crystals. These changes have allowed Davis et al. to quantify the polymorphs of glycine with NIRS during wet granulation [218]. The crystallisation may also be followed in an earlier step, i.e., the production of the active ingredient. Févotte and et al. have used a fibre probe to follow it by means of NIRS. The influence of several parameters was studied. This has shown the possibility of using this spectroscopy to follow the crystallisation of the active on-line [126].

4.1.5. Coating

One of the last steps of the preparation of the drug may be the coating. It is important to ensure the integrity and the good quality of the coating because it may, *e.g.*, determine the release of the drug or ensure stability. The coating can be made on different drug forms such as granulates or tablets. NIR diffuse reflectance spectrometry with a fibre probe has been used to determine the thickness of film coating on pellets by means of a PLS model. The probe was inserted by a side port of the fluidized bed reactor, and located vertically to the pellet bed [156]. In the case of Pérez-Ramos et al. who dealt with tablets, the probe was placed directly in the coating pan for diffuse reflectance measurement. An univariate model was used, which followed the decrease of a specific band of one compound of the core and the increase of a specific band of the coating [219].

4.1.6. Packaging

The very last step in drug manufacturing is the packaging, and even during this last operation a final control can be performed. NIRS spectroscopy has been showed to allow permit a 100% packaging check of tablets. A PCA model was built to sort up to 12,000 tablets per minute according to their quality. For this the spectra given by a NIR camera are compared to the ones used in the model which were acquired on a Foss 6500 Spectrometer [220].

4.1.7. Biotechnology

In 2002–2003 Arnold et al. explained the acquisition, the calibration, the validation and the implementation of fermentation process controlling and monitoring. It was possible to collect NIR spectra in transmittance or in reflectance mode. A second derivative must be calculated. For the calibration, the MLR, PLS or PCR that were compared to the SEC can be used. To calibrate the model it was recommended to use a second set of samples. For the implementation direct on-line or in-line development is advised because the transfer of at-line analysis to in-line analysis would be a challenge [221,222]. Cimander and Mandenius in 2002 applied this on a fermentation of *Escherichia*

coli to produce antibiotics. Spectra were acquired from 400 to 2500 nm with an immersion probe. A PCA and a PLS were used. Models for biomass, tryptophan, phosphate, glucose and acetate were tested. Models for biomass and tryptophan were validated and applied [223]. Fermentations were used to produce toxins or plasmids by Navratil et al. in 2004 as well. They used *Vibrio cholerae* to fabricate cholera-toxin and to produce plasmids. Spectra were acquired between 400 and 2500 nm with a fiber optic probe. A PLS was used to develop calibrations of biomass, glucose and acetate. They were compared to a SEP. Interference problems that occurred when applying the chosen models in the bioreactor are discussed and compensated. Finally they apply NIR prediction models in production [224]. In 2003, Tamburini et al. tried to monitor the fermentation of *Staphylococcus* and *Lactobacillus*. Spectra were acquired from 700 to 1800 nm with a fiber optic probe. Second derivative spectra were calculated from sample minus water difference spectra. A PLS was applied to develop models for glucose, lactic acid, acetic acid and biomass. Those were used for automatic control [225]. Determination of biomass, glucose, lactic acid and acetic acid during fermentations of *Staphylococcus xylosus* was performed by Tosi and et al. in 2003. Spectra were collected from 700 to 1800 nm with a fiber optic probe. A second derivative was calculated and then submitted to a sample selection to discriminate the outliers. Models were developed by PLS for glucose, biomass, lactic acid and acetic acid. The SEC and SEP of models on validation were satisfactory. The model was applied to other microorganisms in the same medium [226]. The study of Yeung et al. compares two strategies for the preparation of calibration samples of a *Saccharomyces cerevisiae* bioprocess. Spectra were obtained between 1900 and 2500 nm. A PLS was applied on the model that was developed for cell debris, protein and RNA. They were compared to the SEP. The selected calibration models were validated [227].

Traditionally many fermentation products come from microbial bioprocesses, however lately mammalian and insect cell cultivations have also been exploited for the high-cost products they can be engineered to produce. Thus Arnold et al. develop a monitoring of mammalian cell cultivation in 2003. Spectra were acquired between 400 and 2500 nm with an immersion probe. A second derivative was calculated and a SNV (standard normal variate) was applied. Models were constructed for glucose, lactate, glutamine and ammonia and compared to the SEC. External and internal (compared to SEP) validation were used [228]. The monitoring for insect cell culture was developed by Riley et al. in 1996. Spectra were acquired between 2000 and 2500 nm. Calibration models were established for glutamine and glucose with a PLS and compared to the SEC and SEP. Models are used but only for high concentrations [229]. Studies of cell culture media were made by Jung et al. and Lewis et al. in 2002 and 2000. They measured their spectra between 2000 and 2500 nm with a fiber optic probe. In the first study, the system was coupled at a lab-system to provide a real-time spectral background reference. A smoothing and a PLS was applied to develop the calibration model of glucose and lactate. Models were compared to SEC, SEP and mean percent error (MPE) [230]. In the second study, they developed models with MSC and without MSC (raw spectra) to predict glucose using PLS. Results were compared to SEC, SEP and MPE and the better model to control the cell culture was used [231]. Results prove the ability of NIR spectroscopy to monitor on-line fermentations and cell cultures.

4.2. Calibration transfer issues

The calibration development is time consuming and expensive due to the reference analyses. Therefore, a new calibration cannot be reproduced for each new spectrometer.

The calibration transfer aim is to apply a unique model on several spectrometers with equivalent prediction errors. The model is developed on a first spectrometer called "master" instrument and is used on several other spectrometers, i.e., "slave" devices. However, even instruments produced by one manufacturer do not produce exactly the same spectra. Without calibration transfer, the slave instrument standard error of prediction will be higher than the one of the master system [17].

Thus solutions to transfer the calibration between devices have to be found. Reviews concerning the calibration transfer methods was written by Fearn [232] and by Feudaleet al. [233]. Three different approaches are available to solve the calibration transfer problem:

The first method is the construction of a robust model which gives us an accurate prediction on several instruments. The calibrations are developed with spectra acquired on several spectrometers. The idea is to include all the variation sources in the calibration. Spectral pretreatment is also useful to reduce the differences between instruments and to build a robust model.

The second one is based on the predicted value correction. A bias or/and a slope between the predicted values of the master and the slave spectrometers is computed. A F-test determines if the slope correction is really needed. Most of the time only a bias correction is enough to correct the differences between instruments. The bias formula is as follows: $\text{bias} = \sum_{j=1}^{n}(y_{\text{master},j} - y_{\text{slave},j})/n$ with y_{slave} are the values predicted with the slave instrument and y_{master} the values predicted with the master instrument and n is the number of samples in the calibration transfer dataset.

The equation $y_{\text{slave}} = y_{\text{master}} - \text{bias}$ is computed with the standardization set. The bias is then applied for each new prediction. The disadvantage is that this method corrects only the systematic error. Thus, the random error on the slave instrument will be unchanged.

The third method is a spectral correction. The slave and master spectra are compared. Then a mathematical correction is applied on the slave spectra in order to make them similar to the master ones and thus to use the master calibration on the slave instruments.

The three most common methods for spectral correction are the direct standardization (DS), the piecewise direct standardization (PDS), and the Shenk and Westerhaus method (SW). DS and PDS methods perform PLS regressions between the slave and the master spectra. DS works with the full spectra whereas PDS performs PLS regressions on moving spectral windows. The SW method [234,235] proceeds in two steps: the first one is a wavelength adjustment between master and slave and the

second one is an absorbance correction with a linear regression for each wavelength. Several other chemometrics methods can be applied: Target Factor Analysis [236], neural networks [237], Wavelets [238], Orthogonal Signal Correction [239]. However, these methods are less applied.

Several samples can be used to perform a spectral correction: stable standard (such as rare earth oxides), pure chemical products or pharmaceutical product subsets. The standards should be stable and representative. It was shown that the calibration transfer is more accurate when the calibration samples and the transfer samples have similar spectra.

Concerning the practical aspect, we have to mention that the calibration transfer procedure has to be validated with an external dataset. Several studies compared the transfer methods for pharmaceutical applications [240–242]. However, no general rules can be proposed and the calibration transfer solution depends on the application and the devices.

5. Conclusion

The potential power of NIRS in quality control and process analytics needs no further demonstration. NIRS is a powerful way to discriminate pharmaceutical compounds. This method can be used qualitatively to detect, to identify, and to qualify raw materials and to control final products. NIRS is a suitable method for classification but also for quantification of pharmaceutics. It is a useful tool for quality control and on-line applications. Near infrared spectroscopy is a potentially precious diagnostic tool in process trouble-shooting and can provide a fingerprinting of pharmaceutical products.

The success of this analytical technique relies on ground advantages [60]. As seen before, NIRS involves the chemical and the physical properties of the samples. This spectroscopy requires no or reduced sample preparation and is non-destructive. Moreover the measurement is fast, e.g., it can be performed in less than a second for on-line applications. NIR frequencies go through glass.

Finally, within the last years NIR imaging systems were developed. A hyper-spectral imaging spectrometer records simultaneously spectral and spatial information of samples. NIR imaging [243] complements NIR spectroscopy and is used when spatial distribution is an important issue of the analysis. This method will be an useful tool in the future.

References

[1] K.H. Norris, J. Infrared Spectrosc. 4 (1996) 31–37.
[2] D.A. Burns, E.W. Ciurczak (Eds.), Handbook of Near-Infrared Analysis, Revised and Expanded, 2nd ed., Marcel Dekker, Inc., New York, Basel, Hingkong, 2001.
[3] H.W. Siesler, Y. Ozaki, S. Kawata, H.M. Heise, Near-Infrared Spectroscopy Principles Instruments Applications, Wiley-VCH, Weinheim, 2002.
[4] L. Bokobza, J. Infrared Spectrosc. 6 (1998) 3–17.
[5] Y. Roggo, L. Duponchel, J.-P. Huvenne, J. Agric. Food Chem. 52 (2004) 1051–1061.
[6] P. Williams, K. Norris, Near-Infrared Technology in the Agricultural and Food Industries, USA, 1987.
[7] M.S. Larrechi, M.P. Callao, TrAC, Trends Anal. Chem. 22 (2003) 634–640.
[8] M. Blanco, S. Maspoch, I. Villarroya, X. Peralta, J.M. Gonzalez, J. Torres, Appl. Spectrosc. 55 (2001) 834–839.
[9] L. Breiman, R. Friedman, R. Olsen, C. Stone, Classification and Regression Trees, Wadsworth, Pacific Grove, CA, 1984.
[10] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part A, Elsevier, Amsterdam, 1997.
[11] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S.D. Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part B, Elsevier, Amsterdam, 1998.
[12] Lavine, Anal. Chem. 72 (2000).
[13] M. Otto, Chemometrics Statistics and Computer Application in Analytical Chemistry, Wiley-VCH, Weinheim, 1999.
[14] R.G. Brereton, Chemometrics Data Analysis for the Laboratory and Chemical Plant, John Wiley & Sons, Chichester, 2003.
[15] D.L. Massart, B.G.M. Vandeginste, S.M. Deming, Y. Michotte, L. Kaufmann, Chemometrics:a textbook, Elsevier, Amsterdam, 2003.
[16] H. Martens, T. Naes, Multivariate Calibration, John Wiley & Sons, Chichester, 1996.
[17] E.W. Ciurczak, J.K. Drennen III (Eds.), Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy, Marcel Dekker, New York, 2002.
[18] Blanco, Coello, Iturriaga, Maspoch, de la Pezuela, Analyst 123 (1998) 135R–150R.
[19] G. Reich, Adv. Drug Deliv. Rev. 57 (2005) 1109–1143.
[20] J. Luypaert, D.L. Massart, Y. Vander Heyden, Talanta 72 (2007) 865–883.
[21] D. Gonzalez-Arjona, V. Gonzalez-Gallero, F. Pablos, A. Gustavo Gonzalez, Anal. Chim. Acta 381 (1999) 257–264.
[22] L. Simon, M. Nazmul Karim, Biochem. Eng. J 7 (2001) 41–48.
[23] F. Marini, F. Balestrieri, R. Bucci, A.D. Magri, A.L. Magri, D. Marini, Chemom. Intell. Lab. Syst. 73 (2004) 85–93.
[24] H. Martens, M. Martens, Multivariate Analysis of Quality An Introduction, John Wiley & Sons, Chichester, 2001.
[25] Y. Roggo, C. Roeseler, M. Ulmschneider, J. Pharm. Biomed. Anal. 36 (2004) 777–786.
[26] W. Vos L. Evers, http://www.stats.ox.ac.uk/∼vos/FullNotes.pdf, 2004.
[27] The European Agency for the Evaluation of Medicinal Products, EMEA/CVMP/961/01, 2003.
[28] M. Daszykowski, B. Walczak, D.L. Massart, Chemom. Intell. Lab. Syst. 56 (2001) 83–92.
[29] M. Daszykowski, B. Walczak, D.L. Massart, Anal. Chim. Acta 468 (2002) 91–103.
[30] J.A. Lopes, P.F. Costa, T.P. Alves, J.C. Menezes, Chemom. Intell. Lab. Syst. 74 (2004) 269–275.
[31] K. Danzer, H. Hobert, C. Fischbacher, K.-U. Jagemann, Chemometrik Grundlagen und Anwendungen, Springer, Berlin, 2001.
[32] T. Mark, Anal. Chem. 57 (1985) 1449–1456.
[33] D. Coomans, M. Jonckheer, D.L. Massart, I. Broeckaert, P. Blockx, Anal. Chim. Acta 103 (1978) 409–415.
[34] M. Sharaf, D. Illman, B. Kowalski, Chemometrics, Wiley/Interscience, New York, 1986.
[35] T. Cover, P. Hart, IEEE T. Inf. Theory 13 (1967) 21–27.
[36] M.-P. Derde, L. Buydens, C. Guns, D.L. Massart, P.K. Hopke, Anal. Chem. 59 (1987) 1868–1871.
[37] D. Coomans, D.L. Massart, Anal. Chim. Acta 138 (1982) 153–165.
[38] A. Candolfi, W. Wu, D.L. Massart, S. Heuerding, J. Pharm. Biomed. Anal. 16 (1998) 1329–1347.
[39] W. Wu, Y. Mallet, B. Walczak, W. Penninckx, D.L. Massart, S. Heuerding, F. Erni, Anal. Chim. Acta 329 (1996) 257–265.
[40] S. Wold, Pattern Recogn. 8 (1976) 127–139.
[41] A. Candolfi, R. De Maesschalck, D.L. Massart, P.A. Hailey, A.C.E. Harrington, J. Pharm. Biomed. Anal. 19 (1999) 923–935.
[42] L. Stahle, S. Wold, J. Chemom. 1 (1987) 185–194.
[43] M. Tenenhaus, La regression PLS Théorie et pratique, Technip, Paris, 1998.

[44] E. Atti, S. Gomez, S.M. Wahl, R. Mendelsohn, E. Paschalis, A.L. Boskey, Bone 31 (2002) 675–684.
[45] R. De Maesschalck, T. Van den Kerkhof, J. Pharm. Biomed. Anal. 37 (2005) 109–114.
[46] R. Leardi, Nature-inspired Methods in Chemometrics: Genetic Algorithms and Artificial Neural Networks, Elsevier, Amsterdam, 2003.
[47] S. Haykin, Neural Networks a Comprehensive Foundation, Prentice Hall, Upper Saddle River, 1999.
[48] R.E. Shaffer, S.L. Rose-Pehrsson, R.A. McGill, Anal. Chim. Acta 384 (1999) 305–317.
[49] H. Wang, D. Chen, Y. Chen, Chemom. Intell. Lab. Syst. 70 (2004) 23–31.
[50] D.F. Specht, Neural Networks 3 (1990) 109–118.
[51] T. Kohonen, Proc. IEEE 78 (1990) 1464–1480.
[52] X. Cui, Z. Zhang, Y. Ren, S. Liu, P.d.B. Harrington, Talanta 64 (2004) 943–948.
[53] R. Burbidge, M. Trotter, B. Buxton, S. Holden, Comput. Chem. Eng. 26 (2001) 5–14.
[54] E.S. Huang, Drug Discov. Today 10 (2005) 69–73.
[55] International Conference on Harmonisation Q2A, 60, 1995, 11260.
[56] International Conference on Harmonisation Q2B, 62, 1997, 27463–27467.
[57] International Conference on Harmonisation Q6A, 65, 2000, 83041–83063.
[58] International Conference on Harmonisation Q7A, 66, 2001, 49028–49029.
[59] Council of Europe (COE). European Pharmacopoeia, 6th ed. European Directorate for the Quality of Medicines. Council of Europe, 2006. ISBN 10: 9287160546.
[60] C. van der Vlies, Eur. J. Pharm. Sci. 2 (1994) 79–81.
[61] M. Ulmschneider, G. Barth, E. Trenka, Pharmazeutische Industrie 62 (2000) 374–376.
[62] M. Ulmschneider, E. Penigault, Analusis 28 (2000) 83–87.
[63] M. Ulmschneider, E. Penigault, Analusis 28 (2000) 336–346.
[64] M. Ulmschneider, A. Wunenburger, E. Penigault, Analusis 27 (1999) 854–856.
[65] N.K. Ebube, S.S. Thosar, R.A. Roberts, M.S. Kemper, R. Rubinovitz, D.L. Martin, G.E. Reier, T.A. Wheatley, A.J. Shukla, Pharm. Dev. Technol. 4 (1999) 19–26.
[66] K. Kramer, S. Ebel, Anal. Chim. Acta 420 (2000) 155–161.
[67] K. Kreft, B. Kozamernik, U. Urleb, Int. J. Pharm. 177 (1999) 1–6.
[68] M. Andre, Anal. Chem. 75 (2003) 3460–3467.
[69] W.L. Yoon, R.D. Jee, A.C. Moffat, P.D. Blackler, K. Yeung, D.C. Lee, Analyst 124 (1999) 1197–1203.
[70] C.I. Gerhausser, K.-A. Kovar, Appl. Spectrosc. 51 (1997) 1504–1510.
[71] N. Sondermann, K.-A. Kovar, Foren. Sci. Int. 102 (1999) 133–147.
[72] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, C. Perez-Maseda, Anal. Chim. Acta 407 (2000) 247–254.
[73] K.L. Vora, G. Buckton, D. Clapham, Eur. J. Pharm. Sci. 22 (2004) 97–105.
[74] M. Blanco, D. Valdes, M.S. Bayod, F. Fernandez-Mari, I. Llorente, Anal. Chim. Acta 502 (2004) 221–227.
[75] G. Buckton, E. Yonemochi, J. Hammond, A. Moffat, Int. J. Pharm. 168 (1998) 231–241.
[76] W.L. Yoon, R.D. Jee, A. Charvill, G. Lee, A.C. Moffat, J. Pharm. Biomed. Anal. 34 (2004) 933–944.
[77] S.H.F. Scafi, C. Pasquini, Analyst 126 (2001) 2218–2224.
[78] M.J. Vredenbregt, L. Blok-Tip, R. Hoogerbrugge, D.M. Barends, D.d. Kaste, J. Pharm. Biomed. Anal. 40 (2006) 840–849.
[79] B.J. Westenberger, C.D. Ellison, A.S. Fussner, S. Jenney, R.E. Kolinski, T.G. Lipe, R.C. Lyon, T.W. Moore, L.K. Revelle, A.P. Smith, Int. J. Pharm. 306 (2005) 56–70.
[80] O.Y. Rodionova, L.P. Houmoller, A.L. Pomerantsev, P. Geladi, J. Burger, V.L. Dorofeyev, A.P. Arzamastsev, Anal. Chim. Acta 549 (2005) 151–158.
[81] Y. Roggo, N. Jent, A. Edmond, P. Chalus, M. Ulmschneider, Eur. J. Pharm. Biopharm. 61 (2005) 100–110.
[82] M. Blanco, R. Gozalez Bano, E. Bertran, Talanta 56 (2002) 203–212.
[83] L.E. Rodriguez-Saona, F.M. Khambaty, F.S. Fry, E.M. Calvey, J. Agric. Food Chem. 49 (2001) 574–579.
[84] L.E. Rodriguez-Saona, F.M. Khambaty, F.S. Fry, J. Dubois, E.M. Calvey, J. Food Protect. 67 (2004) 2555–2559.
[85] W.L. Yoon, A.C. Moffat, Analyst 123 (1998) 1029–1034.
[86] W.L. Yoon, N.C. North, R.D. Jee, A.C. Moffat, Proceedings of the 9th International Conference on Near Infrared Spectroscopy, Verona, Italy, June 13–18, 1999, pp. 547–550.
[87] R.J. Barnes, M.S. Dhanoa, S.J. Lister, J. Infrared Spectrosc. 1 (1993) 185–186.
[88] M.S. Dhanoa, S.J. Lister, R. Sanderson, R.J. Barnes, J. Infrared Spectrosc. 2 (1994) 43–47.
[89] P.J. de Groot, G.J. Postma, W.J. Melssen, L.M.C. Buydens, Appl. Spectrosc. 55 (2001) 173–181.
[90] W. Wu, B. Walczak, D.L. Massart, K.A. Prebble, I.R. Last, Anal. Chim. Acta 315 (1995) 243–255.
[91] A. Candolfi, R. De Maesschalck, D. Jouan-Rimbaud, P.A. Hailey, D.L. Massart, J. Pharm. Biomed. Anal. 21 (1999) 115–132.
[92] M. Blanco, M.A. Romero, Analyst 126 (2001) 2212–2217.
[93] A. Candolfi, D.L. Massart, Appl. Spectrosc. 54 (2000) 48–53.
[94] W.L. Yoon, R.D. Jee, A.C. Moffat, Am. Pharm. Rev. 5 (2002) 106–111.
[95] M.J. Vredenbregt, P.W.J. Caspers, R. Hoogerbrugge, D.M. Barends, Eur. J. Pharm. Biopharm. 56 (2003) 489–499.
[96] T. Naes, T. Isaksson, T. Fearn, A.M.C. Davies, A User-Friendly Guide to Multivariate Calibration and Classification, NIR Publications, Chichester, UK, 2002.
[97] M.J. Adams, Chemometrics in Analytical Spectroscopy, Royal Society of Chemistry, Cambridge, UK, 1995.
[98] S. Wold, J. Trygg, A. Berglund, H. Antti, Chemom. Intell. Lab. Syst. 58 (2001) 131–150.
[99] T. Fearn, Chemom. Intell. Lab. Syst. 50 (2000) 47–52.
[100] C.K. Markopoulou, E.T. Malliou, J.E. Koundourellis, J. Pharm. Biomed. Anal. 37 (2005) 249–258.
[101] J.-H. Jiang, R.J. Berry, H.W. Siesler, Y. Ozaki, Anal. Chem. 74 (2002) 3555–3565.
[102] Y. Dou, Y. Sun, Y. Ren, P. Ju, Y. Ren, J. Pharm. Biomed. Anal. 37 (2005) 543–549.
[103] Y. Chen, Drug Dev. Ind. Pharm. 27 (2001) 623–631.
[104] A.P. Plumb, R.C. Rowe, P. York, C. Doherty, Eur. J. Pharm. Sci. 16 (2002) 281–288.
[105] A.P. Plumb, R.C. Rowe, P. York, M. Brown, Eur. J. Pharm. Sci. 25 (2005) 395–405.
[106] U. Thissen, M. Pepers, B. Ustun, W.J. Melssen, L.M.C. Buydens, Chemom. Intell. Lab. Syst. 73 (2004) 169–179.
[107] U. Thissen, B. Ustun, W.J. Melssen, L.M.C. Buydens, Anal. Chem. 76 (2004) 3099–3105.
[108] F. Chauchard, R. Cogdill, S. Roussel, J.M. Roger, V. Bellon-Maurel, Chemom. Intell. Lab. Syst. 71 (2004) 141–150.
[109] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, J. Pages, Anal. Chim. Acta 384 (1999) 207–214.
[110] E. Pretsch, C.L. Wilkins, TrAC, Trends Anal. Chem. 25 (2006) 1043–1138.
[111] M.P. Freitas, A. Sabadin, L.M. Silva, F.M. Giannotti, D.A. do Couto, E. Tonhi, R.S. Medeiros, G.L. Coco, V.F.T. Russo, Martins, et al., J. Pharm. Biomed. Anal. 39 (2005) 17–21.
[112] K.M. Morisseau, C.T. Rhodes, Pharm. Res. 14 (1997) 108–111.
[113] J.D. Kirsch, J.K. Drennen, J. Pharm. Biomed. Anal. 19 (1999) 351–362.
[114] J.-H. Guo, G.W. Skinner, W.W. Harcum, J.P. Malone, L.G. Weyer, Drug Dev. Ind. Pharm. 25 (1999) 1267–1270.
[115] M. Blanco, M. Alcala, Anal. Chim. Acta 557 (2006) 353–359.
[116] M. Donoso, E.S. Ghaly, Pharm. Dev. Technol. 9 (2004) 247.
[117] O. Berntsson, L.G. Danielsson, S. Folestad, Anal. Chim. Acta 364 (1998) 243–251.
[118] O. Berntsson, T. Burger, S. Folestad, L.G. Danielsson, J. Kuhn, J. Fricke, Anal. Chem. 71 (1999) 617–623.
[119] M. Otsuka, Powder Technol. 141 (2004) 244–250.
[120] I. Fix, K.-J. Steffens, Drug Dev. Ind. Pharm. 30 (2004) 513–523.
[121] A. Gombas, I. Antal, P. Szabo-Revesz, S. Marton, I. Eros, Int. J. Pharm. 256 (2003) 25–32.

[122] W. Li, G.D. Worosila, W. Wang, T. Mascaro, J. Pharm. Sci. 94 (2005) 2800–2806.
[123] S.J. Bai, M. Rani, R. Suryanarayanan, J.F. Carpenter, R. Nayar, M.C. Manning, J. Pharm. Sci. 93 (2004) 2439–2447.
[124] J.J. Seyer, P.E. Luner, Pharm. Dev. Technol. 6 (2001) 573–582.
[125] O. Berntsson, L.G. Danielsson, M.O. Johansson, S. Folestad, Anal. Chim. Acta 419 (2000) 45–54.
[126] G. Fevotte, J. Calas, F. Puel, C. Hoff, Int. J. Pharm. 273 (2004) 159–169.
[127] A.D. Patel, P.E. Luner, M.S. Kemper, Int. J. Pharm. 206 (2000) 63–74.
[128] M. Savolainen, A. Heinz, C. Strachan, K.C. Gordon, J. Yliruusi, T. Rades, N. Sandler, Eur. J. Pharm. Sci., in press, Corrected Proof.
[129] W.P. Findlay, G.R. Peck, K.R. Morris, J. Pharm. Sci. 94 (2005) 604–612.
[130] X. Zhou, P. Hines, M.W. Borer, J. Pharm. Biomed. Anal. 17 (1998) 219–225.
[131] J. Rantanen, O. Antikainen, J.-P. Mannermaa, J. Yliruusi, Pharm. Dev. Technol. 5 (2000) 209.
[132] R.G. Buice Jr., T.B. Gold, R.A. Lodder, G.A. Digenis, Pharm. Res. 12 (1995) 161–163.
[133] O. Berntsson, G. Zackrisson, G. Ostling, J. Pharm. Biomed. Anal. 15 (1997) 895–900.
[134] N.W. Broad, R.D. Jee, A.C. Moffat, M.J. Eaves, W.C. Mann, W. Dziki, Analyst 125 (2000) 2054–2058.
[135] H. Yang, J. Irudayaraj, J. Pharm. Pharmacol. 54 (2002) 1247–1255.
[136] M. Iyer, H.R. Morris, J.K. Drennen III, J. Infrared Spectrosc. 10 (2002) 233–245.
[137] P. Merckle, K.A. Kovar, J. Pharm. Biomed. Anal. 17 (1998) 365–374.
[138] R.C. Schneider, K.-A. Kovar, Foren. Sci. Int. 134 (2003) 187–195.
[139] S.S. Thosar, Pharm. Dev. Technol. 6 (2001) 19–29.
[140] P. Chalus, Y. Roggo, S. Walter, M. Ulmschneider, Talanta 66 (2005) 1294–1302.
[141] P. Chalus, Y. Roggo, S. Freitag, M. Ulmschneider, Spectra Anal. (2005) 44–49.
[142] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, D. Serrano, Analyst 123 (1998) 2307–2312.
[143] S.M. Han, P.G. Faulkner, J. Pharm. Biomed. Anal. 14 (1996) 1681–1689.
[144] I. Rager, G. Roos, P.C. Schmidt, K.A. Kovar, J. Pharm. Biomed. Anal. 28 (2002) 439–446.
[145] M. Kemper, E. Magnuson, S. Lowry, W. McCarthy, N. Aksornkoae, D. Watts, J. Johnson, A. Shukla, AAPS Pharmsci. 3 (2001).
[146] S.C. Baratieri, J.M. Barbosa, M.P. Freitas, J.A. Martins, J. Pharm. Biomed. Anal. 40 (2006) 51–55.
[147] W. Fountain, K. Dumstorf, A.E. Lowell, R.A. Lodder, R.J. Mumper, J. Pharm. Biomed. Anal. 33 (2003) 181–189.
[148] G.K. Webster, D.A. Farrand, E. Johnson, M.A. Litchman, N. Broad, S. Maris, J. Pharm. Biomed. Anal. 33 (2003) 21–32.
[149] N.D. Wilson, M.S. Ivanova, R.A. Watt, A.C. Moffat, J. Pharm. Pharmacol. 54 (2002) 1257–1263.
[150] N.D. Wilson, R.A. Watt, A.C. Moffat, J. Pharm. Pharmacol. 53 (2001) 95–102.
[151] R.A. Forbes, M.Z. Luo, D.R. Smith, J. Pharm. Biomed. Anal. 25 (2001) 239–256.
[152] C.C. Fertig, F. Podczeck, R.D. Jee, M.R. Smith, Eur. J. Pharm. Sci. 21 (2004) 155–159.
[153] S.V.S. Tumuluri, S. Prodduturi, M.M. Crowley, S.P. Stodghill, J.W. McGinity, M.A. Repka, B.A. Avery, Drug Dev. Ind. Pharm. 30 (2004) 505–511.
[154] G.X. Zhou, Z. Ge, J. Dorwart, B. Izzo, J. Kukura, G. Bicker, J. Wyvratt, J. Pharm. Sci. 92 (2003) 1058–1065.
[155] M. Gaub, C. Roeseler, G. Roos, K.-A. Kovar, J. Pharm. Biomed. Anal. 36 (2004) 859–864.
[156] M. Andersson, S. Folestad, J. Gottfries, M.O. Johansson, M. Josefson, K.G. Wahlund, Anal. Chem. 72 (2000) 2099–2108.
[157] M. Andersson, M. Josefson, F.W. Langkilde, K.G. Wahlund, J. Pharm. Biomed. Anal. 20 (1999) 27–37.
[158] I.H.I. Habib, M.S. Kamel, Talanta 60 (2003) 185–190.
[159] M. Laasonen, T. Harmia-Pulkkinen, C. Simard, M. Raesaenen, H. Vuorela, Anal. Chem. 75 (2003) 754–760.
[160] N.W. Broad, R.D. Jee, A.C. Moffat, M.R. Smith, Analyst 126 (2001) 2207–2211.
[161] M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, N. Pou, Fresen. J. Anal. Chem. 368 (2000) 534–539.
[162] J.L. Ramirez, M.K. Bellamy, R.J. Romanach, AAPS Pharmsci. 2 (2001) E11–E14.
[163] A. Eustaquio, M. Blanco, R.D. Jee, A.C. Moffat, Anal. Chim. Acta 383 (1999) 283–290.
[164] A.D. Trafford, R.D. Jee, A.C. Moffat, P. Graham, Analyst 124 (1999) 163–167.
[165] Y. Dou, Y. Sun, Y. Ren, Y. Ren, Anal. Chim. Acta 528 (2005) 55–61.
[166] M. Blanco, J. Coello, A. Eustaquio, H. Iturriaga, S. Maspoch, J. Pharm. Sci. 88 (1999) 551–556.
[167] P. Corti, E. Ceramelli, G. Dreassi, S. Mattii, Analyst 124 (1999) 755–758.
[168] M. Blanco, M.A. Romero, J. Pharm. Biomed. Anal. 30 (2002) 467–472.
[169] Y.-C. Feng, C.-Q. Hu, J. Pharm. Biomed. Anal. 41 (2006) 373–384.
[170] I.R. Last, K.A. Prebble, J. Pharm. Biomed. Anal. 11 (1993) 1071–1076.
[171] L. Kamat, DeLuca, Pharm. Res. 6 (1989) 961–965.
[172] J.A. Jones, I.R. Last, B.F. MacDonald, K.A. Prebble, J. Pharm. Biomed. Anal. 11 (1993) 1227–1231.
[173] M. Savage, J. Torres, L. Franks, B. Masecar, J. Hotta, Biologicals 26 (1998) 119–124.
[174] T.P. Lin, C.C. Hsu, Pharm. Sci. Technol. (2001) 196–204.
[175] Cao, Mao, Chen, Lin, Krishnan, Cauchon, J. Pharm. Sci. 95 (2006) 2077–2086.
[176] A. Stokvold, K. Dyrstad, F.O. Libnau, J. Pharm. Biomed. Anal. 28 (2002) 867–873.
[177] M.W.J. Derksen, P.J.M. van de Oetelaar, F.A. Maris, J. Pharm. Biomed. Anal. 17 (1998) 473–480.
[178] S. Bai, R. Nayar, J.F. Carpenter, M.C. Manning, J. Pharm. Sci. 94 (2005) 2030–2038.
[179] K.-I. Izutsu, Y. Fujimaki, A. Kuwabara, Y. Hiyama, C. Yomota, N. Aoyagi, J. Pharm. Sci. 95 (2006) 781–789.
[180] K.-I. Izutsu, Y. Fujimaki, A. Kuwabara, N. Aoyagi, Int. J. Pharm. 301 (2005) 161–169.
[181] J. Liu, Pharm. Dev. Technol. 11 (2006) 3–28.
[182] L. Sukowski, M. Ulmschneider, Pharm. Ind. 67 (2005) 830–835.
[183] Brülls, Folestad, Spàren, Rasmuson, Pharmaceut. Res. (2003) 494–499.
[184] S.A. Arnold, J. Crowley, S. Vaidyanathan, L. Matheson, P. Mohan, J.W. Hall, L.M. Harvey, B. McNeil, Enzyme Microb. Technol. 27 (2000) 691–697.
[185] S. Vaidyanathan, A. Arnold, L. Matheson, P. Mohan, G. Macaloney, B. McNeil, L.M. Harvey, Biotechnol. Prog. 16 (2000) 1098–1105.
[186] S. Vaidyanathan, L.M. Harvey, B. McNeil, Anal. Chim. Acta 428 (2001) 41–59.
[187] S. Vaidyanathan, G. Macaloney, L.M. Harvey, B. McNeil, Appl. Spectrosc. 55 (2001) 444–453.
[188] B.L. Frazier, P. Larmour, M.R. Riley, Biotechnol. Bioeng. 72 (2000) 364–368.
[189] M. Rhiel, M.B. Cohen, D.W. Murhammer, M.A. Arnold, Biotechnol. Bioeng. 77 (2001) 73–82.
[190] M.J. McShane, G.L. Coté, Appl. Spectrosc. 52 (1998) 1073–1078.
[191] M. Laasonen, T. Harmia-Pulkkinen, C. Simard, M. Rasanen, H. Vuorela, Eur. J. Pharm. Sci. 21 (2004) 493–500.
[192] M. Blanco, M.A. Romero, M. Alcala, Talanta 64 (2004) 597–602.
[193] Blanco, Coello, Iturriaga, Maspoch, de la Pezuela, Analyst 122 (1997) 761–765.
[194] J.H. Kalivas, Anal. Chim. Acta 428 (2001) 31–40.
[195] H.A.D. Filho, R.K.H. Galvao, M.C.U. Araujo, E.C. da Silva, T.C.B. Saldanha, G.E. Jose, C. Pasquini, J. Milton Raimundo, J.J.R. Ivo, Rohwedder, Chemom. Intell. Lab. Syst. 72 (2004) 83–91.
[196] R.K.H. Galvao, M.C.U. Araujo, G.E. Jose, M.J.C. Pontes, E.C. Silva, T.C.B. Saldanha, Talanta 67 (2005) 736–740.
[197] W. Wu, B. Walczak, D.L. Massart, S. Heuerding, F. Erni, I.R. Last, K.A. Prebble, Chemom. Intell. Lab. Syst. 33 (1996) 35–46.
[198] T. Naes, T. Isaksson, T. Fearn, T. Davies, Multivariate Calibration and Classification, NIR Publications, Chichester, UK, 2002.

[199] D.J. Wargo, J.K. Drennen, J. Pharm. Biomed. Anal. 14 (1996) 1415–1423.
[200] R.P. Cogdill, C.A. Anderson, M. Delgado-Lopez, D. Molseed, R. Chisholm, R. Bolton, T. Herkert, A.M. Afnan, J.K.E.A. Drennen, AAPS Pharmsci. 6 (2005) E262–E272.
[201] R.P. Cogdill, C.A. Anderson, M. Delgado, R. Chisholm, R. Bolton, T. Herkert, A.M. Afnan, J.K. Drennen III, AAPS Pharmsci. 6 (2005) E273–E283.
[202] R.P. Cogdill, C.A. Anderson, J.K. Drennen, AAPS Pharmsci. 6 (2005).
[203] J. Cho, P.J. Gemperline, P.K. Aldridge, S.S. Sekulic, Anal. Chim. Acta 348 (1997) 303–310.
[204] P.A. Hailey, P. Doherty, P. Tapsell, T. Oliver, P.K. Aldridge, J. Pharm. Biomed. Anal. 14 (1996) 551–559.
[205] S.S. Sekulic, H.W. Ward II, D.R. Brannegan, E.D. Stanley, C.L. Evans, S.T. Sciavolino, P.A. Hailey, P.K. Aldridge, Anal. Chem. 68 (1996) 509–513.
[206] A.S. El-Hagrasy, H.R. Morris, F. D'Amico, R.A. Lodder, J.K. Drennen, J. Pharm. Sci. 90 (2001) 1298–1307.
[207] S.S. Sekulic, J. Wakeman, P. Doherty, P.A. Hailey, J. Pharm. Biomed. Anal. 17 (1998) 1285–1309.
[208] E.T.S. Skibsted, H.F.M. Boelens, J.A. Westerhuis, D.T. Witte, A.K. Smilde, J. Pharm. Biomed. Anal. 41 (2006) 26–35.
[209] E.T.S. Skibsted, J.A. Westerhuis, A.K. Smilde, D.T. Witte, J. Pharm. Biomed. Anal. 43 (2007) 1297–1305.
[210] M. Popo, S. Romero-Torres, C. Conde, R.J. Romanach, AAPS Pharmsci. 3 (2002) E24.
[211] O. Berntsson, L.-G. Danielsson, B. Lagerholm, S. Folestad, Powder Technol. 123 (2002) 185–193.
[212] J. Rantanen, S. Lehtola, P. Ramet, J.-P. Mannermaa, J. Yliruusi, Powder Technol. 99 (1998) 163–170.
[213] J. Rantanen, E. Rasanen, J. Tenhunen, M. Kansakoski, J.-P. Mannermaa, J. Yliruusi, Eur. J. Pharm. Biopharm. 50 (2000) 271–276.
[214] J. Rantanen, E. Rasanen, O. Antikainen, J.-P. Mannermaa, J. Yliruusi, Chemom. Intell. Lab. Syst. 56 (2001) 51–58.
[215] P. Frake, D. Greenhalgh, S.M. Grierson, J.M. Hempenstall, D.R. Rudd, Int. J. Pharm. 151 (1997) 75–80.
[216] A. Gupta, G.E. Peck, R.W. Miller, K.R. Morris, J. Pharm. Sci. 93 (2004) 1047–1053.
[217] A. Gupta, G.E. Peck, R.W. Miller, K.R. Morris, J. Pharm. Sci. 94 (2005) 1589–1597.
[218] T.D. Davis, G.E. Peck, J.G. Stowell, K.R. Morris, S.R. Byrn, Pharm. Res. 21 (2004) 860–866.
[219] J.D. Perez-Ramos, W.P. Findlay, G. Peck, K.R. Morris, AAPS Pharmsci. 6 (2005).
[220] T. Herkert, H. Prinz, K.-A. Kovar, Eur. J. Pharm. Biopharm. 51 (2001) 9–16.
[221] S.A. Arnold, L.M. Harvey, B. McNeil, J.W. Hall, Biopharm. Intern., November 2002 (2002).
[222] S.A. Arnold, L.M. Harvey, B. McNeil, J.W. Hall, Biopharm. Intern. January 2003 (2003).
[223] C. Cimander, C.-F. Mandenius, J. Chem. Technol. Biotechnol. 77 (2002) 1157–1168.
[224] M. Navratil, A. Norberg, L. Lembren, C.-F. Mandenius, J. Biotechnol. 115 (2005) 67–79.
[225] E. Tamburini, G. Vaccari, S. Tosi, A. Trilli, Appl. Spectrosc. 57 (2003) 132–138.
[226] S. Tosi, M. Rossi, E. Tamburini, G. Vaccari, A. Amaretti, D. Matteuzzi, Biotechnol. Prog. 19 (2003) 1816–1821.
[227] K.S.Y. Yeung, M. Hoare, N.F. Thornhill, T. Williams, J.D. Vaghjiani, Biotechnol. Bioeng. 63 (1998).
[228] S. Arnold, J. Crowley, N. Woods, L. Harvey, B. McNeil, Biotechnol. Bioeng. 84 (2003) 13–19.
[229] M.R. Riley, M. Rhiel, X. Zhou, M.A. Arnold, D.W. Murhammer, Biotechnol. Bioeng. 55 (1996) 11–15.
[230] B. Jung, S. Lee, I.H. Yang, T. Good, G.L. Coté, Appl. Spectrosc. 56 (2002) 51–57.
[231] C.B. Lewis, R.J. McNichols, A. Gowda, G.L. Coté, Appl. Spectrosc. 54 (2000) 1453–1457.
[232] T. Fearn, J. Infrared Spectrosc. 9 (2001) 229–244.
[233] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferre, Chemom. Intell. Lab. Syst. 64 (2002) 181–192.
[234] J. Shenk, M. Westeraus, Crop Sci. 31 (1991) 1548–1551.
[235] E. Bouveresse, D.L. Massart, P. Dardenne, Anal. Chim. Acta 297 (1994) 405–416.
[236] Y.-B. Wang, H.-F. Yuan, W.-Z. Lu, Spectrosc. Spectral Anal. 25 (2005) 398–401.
[237] L. Duponchel, C. Ruckebusch, J.-P. Huvenne, P. Legrand, J. Infrared Spectrosc. 7 (1999) 155–166.
[238] J. Yoon, B. Lee, C. Han, Chemom. Intell. Lab. Syst. 64 (2002) 1–14.
[239] J. Sjoblom, O. Svensson, M. Josefson, H. Kullberg, S. Wold, Chemom. Intell. Lab. Syst. 44 (1998) 229–244.
[240] E.-L. Bergman, H. Brage, M. Josefson, O. Svensson, A. Sparen, J. Pharm. Biomed. Anal. 41 (2006) 89–98.
[241] H. Leion, S. Folestad, M. Josefson, A. Sparen, J. Pharm. Biomed. Anal. 37 (2005) 47–55.
[242] M.R. Smith, R.D. Jee, A.C. Moffat, D.R. Rees, N.W. Broad, Analyst 129 (2004) 806–816.
[243] Y. Roggo, A. Edmond, P. Chalus, M. Ulmschneider, Anal. Chim. Acta 535 (2005) 79–87.